DOCUMENT RESUME

ED 202 859 TM 810 178

AUTHOR Clark, Mary Jo
TITLE An Evaluation of GRE Sender Scores as a Measure of Graduate Admissions Selectivity in Two Fields.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RR-81-2
PUB DATE Feb 81
NOTE 18p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Aptitude Tests; Chemistry; *Correlation; *Graduate Study; History; *Program Evaluation; *Scores; *Selective Admission
IDENTIFIERS *Graduate Record Examinations

ABSTRACT

The average Graduate Record Examinations (GRE) Aptitude Test scores of test-takers who had their scores sent to 57 history and 60 chemistry departments in 1972-73 ("scores senders") are compared with the average Aptitude Test scores of first-year students who enrolled in these departments in the fall of 1973 ("enrolled students" or "matriculants"). There is a positive correlation of .5 to .7 between the department mean sender and mean matriculant scores; however, about an equal number of departments in both fields were found to enroll students with average test scores that were relatively lower as well as relatively higher than their average sender scores. Results were similar for departments with high, moderate, and low reputational ratings. Although the self-selection of applicants appears to account for a large part of departmental "selectivity" as this is represented by average GRE Aptitude Test scores, the study concludes that average sender scores should not be used to describe the admissions selectivity of individual graduate departments in these two fields. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED202859

RR-81-2

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

RESEARCH REPORT

AN EVALUATION OF GRE SENDER SCORES AS A MEASURE OF GRADUATE ADMISSIONS SELECTIVITY IN TWO FIELDS

Mary Jo Clark

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. C. Weidenmiller

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 178

Educational Testing Service
Princeton, New Jersey
February 1981

2

Copyright © 1981. Educational Testing Service. All rights reserved.

Abstract

The average GRE Aptitude Test scores of test-takers who had their scores sent to 57 history and 60 chemistry departments in 1972-73 ("score senders") are compared with the average Aptitude Test scores of first-year students who enrolled in these departments in the fall of 1973 ("enrolled students" or "matriculants"). There is a positive correlation of .5 to .7 between the department mean sender and mean matriculant scores; however, about an equal number of departments in both fields were found to enroll students with average test scores that were relatively lower as well as relatively higher than their average sender scores. Results were similar for departments with high, moderate, and low reputational ratings. Though the self-selection of applicants appears to account for a large part of departmental "selectivity" as this is represented by average GRE Aptitude Test scores, the study concludes that average sender scores should not be used to describe the admissions selectivity of individual graduate departments in these two fields.

An Evaluation of GRE Sender Scores as a Measure of Graduate Admissions Selectivity in Two Fields

A regular service of the GRE Program in recent years has been to send universities an annual summary of the GRE test scores that were sent to departments or schools within the university by prospective applicants during the preceding year. In addition to frequency distributions, the report includes test score means and standard deviations whenever 25 or more reports were sent to a given department. Occasionally someone asks whether these mean sender scores might be used as an indirect measure of a given department's level of admissions selectivity. The logic urguing for close correspondence between sender scores and departmental selectivity is based in part on evidence that prospective graduate students tend to apply for admission to departments where they think their chances for admission are good. Therefore, students with higher test scores probably are more likely to apply to selective departments than are students with lower test scores. The hypothesis of correspondence would be supported if the mean GRE scores of senders rank ordered departments in a way that was very similar to a rank ordering based on the mean GRE scores of enrolled students. In this case, the mean sender scores would appear to be a reasonable proxy for the mean matriculant scores.

Though the idea of using mean sender scores as an indication of departmental selectivity seems reasonable, it has never been used in practice because of many unanswered questions about the appropriateness of such use. For example, though it is plausible to assume that most

graduate admissions applicants are fairly realistic in their selection of programs, what proportion of the test score senders never actually file applications for admission to the programs that receive their scores? How many of the applications that are filed represent "safety" choices or "long-shot" applications, rather than realistic matching? Would large numbers of applications to popular or well-known programs result in less representative mean sender scores for these departments than for departments in smaller or regionally oriented universities? Can we really demonstrate a relationship between the level of the test scores that are sent to a department and the average test scores of the applicants who are accepted and who finally enroll in a department? None of these questions could be answered by data routinely available in the GRE program files.

Recently, however, the mean GRE verbal and quantitative aptitude test scores for students enrolling in a sizeable number of graduate history and chemistry departments became available, and provided an opportunity to evaluate mean sender scores in relation to mean enrolled student scores for these departments. The purpose of this report is to present the resulting comparison.

Sender and matriculant mean GRE aptitude test scores were available for 57 history departments and 60 chemistry departments. Departmental sender score means were computed only if at least 25 scores had been sent to the department; mean sender scores for 30 percent of the history departments and 43 percent of the chemistry departments were based on more than 100 score reports. The number of matriculant scores per

program was much smaller. Twenty-one of the history departments (37%) and 24 of the chemistry departments (40%) were represented by fewer than 10 GRE scores for students entering in 1973 (the year in which all of these data were collected). Only four history departments (7%) and six chemistry departments (10%) had GRE scores for 30 or more enrolled students. Some departments admitted students without GRE scores; GRE scores could not be located for others.[1] Despite these problems, these numbers of newly enrolling students appear to reflect the approximate sizes of entering graduate classes in these fields, and to be sufficient for the analyses planned for this study.

The sample also appears to be reasonably representative of doctoral programs in these two fields as guaged by reputational ratings. Based on ratings of the quality of the graduate faculty collected by Roose and Andersen in 1969,[2] six (11%) of the history departments and 12 (20%) of the chemistry departments were rated 3.5 or above ("distinguished" or "strong"), 24 (42%) of the history departments and 34 (57%) of the chemistry departments were rated "good" or "adequate plus" (ratings between 3.49 and 2.0), and 27 (47%) of the history departments and 24 (23%) of the chemistry departments either were not included in the ratings or rated below 2.0 (less than "adequate"). Most programs

---

[1]For details, see Drew, D. E. Science development: An evaluation study. National Board on Graduate Education Technical Report No. 4. Washington, D.C.: National Academy of Sciences, 1975, p. 129.

[2]Roose, K. D., & Andersen, C. J. A rating of graduate programs. Washington, D.C.: American Council on Education, 1970.

without reputational ratings offer only the master's degree, are new doctoral programs, or grant doctor's degrees only occasionally.

Table 1 presents the average departmental mean GRE Aptitude test scores for test senders and matriculants, and the product-moment correlations between the two sets of scores. In general, the average departmental mean scores for enrolled students are higher than the average sender scores, especially verbal scores in history departments, but none of the differences are very large. In fact, in chemistry the average quantitative test results for score senders are almost exactly the same as the average matriculant scores. Correlation coefficients in the .5 to .7 range indicate a definite positive association between the sender and matriculant scores, but also sufficient independence to indicate a number of individual department variations.

When departments with small numbers of matriculant scores, and therefore possibly unstable or unrepresentative mean matriculant scores, are omitted from the calculations, both the average mean scores and the correlations are somewhat larger. Removing small departments has a particularly noticable effect on the correlations between sender and matriculant mean scores in chemistry, suggesting greater sampling error among these departments.

Several additional ways of describing the data were explored in order to examine the association between sender and matriculant scores in more detail. Since most history departments give more weight to verbal scores and most chemistry departments give more weight to quantitative scores in the admissions process, it seemed appropriate to give particular attention to these respective scores in each

Table 1

The Relationship Between Average GRE Aptitude Test Scores of Test Score Senders and Matriculants in a Sample of Graduate Departments

	All Departments		Departments with n ≥ 10	
	V	Q	V	Q
HISTORY				
Number of departments	57	57	36	36
Score Senders				
Mean of dept. means	562	516	569	522
S.D. of dept. means	39	35	43	37
Matriculants				
Mean of dept. means	604	541	612	548
S.D. of dept. means	53	54	49	46
Correlation	.66	.60	.70	.74
CHEMISTRY				
Number of departments	60	60	36	36
Score Senders				
Mean of dept. means	486	660	506	666
S.D. of dept. means	50	24	47	23
Matriculants				
Mean of dept means	514	657	527	669
S.D. of dept. means	60	43	51	33
Correlation	.54	.46	.71	.62

Note: Within-department standard deviations of GRE scores were not available for these analyses. The standard deviation for V and Q scores among all test-takers is about 125. The within-department standard deviation of the GRE scores of enrolled students would be expected to be smaller than this.

department. Results of scatterplots for mean verbal test scores in history departments and mean quantitative test scores in chemistry departments are presented in Figures 1 and 2. Points representing departments with fewer than 10 matriculant scores are circled.

As suggested by the data in Table 1, the mean verbal scores of matriculants in history were generally higher than the mean sender scores as indicated by the fact that most points are above the diagonal line in Figure 1. Departments with a small number of matriculant scores (circled points in the figure) tend to be at the lower end and on the edges of the distribution; in general, however, there are no obvious outliers influencing the size of the correlation coefficients reported in Table 1.

Unlike history, in chemistry departments (Figure 2) about half the departments have mean quantitative sender scores about the same as or higher than matriculant scores (on or below the diagonal line), and there are some outlying pairs of scores that undoubtedly had a depressing effect on the correlations reported in Table 1. Omitting small departments eliminates a number of these outliers, and the correlation increases from .46 to .62. Still, the range of mean quantitative scores in chemistry is fairly narrow (roughly 600 to 720) compared to the range of mean verbal scores in history (520 to 720). The narrow range of quantitative scores acceptable for admission in this field may help to explain why the average GRE scores of senders are often higher than the average GRE scores of enrolled students.

A number of chemistry departments with higher sender than matriculant GRE scores may also reflect a higher level of multiple applications than

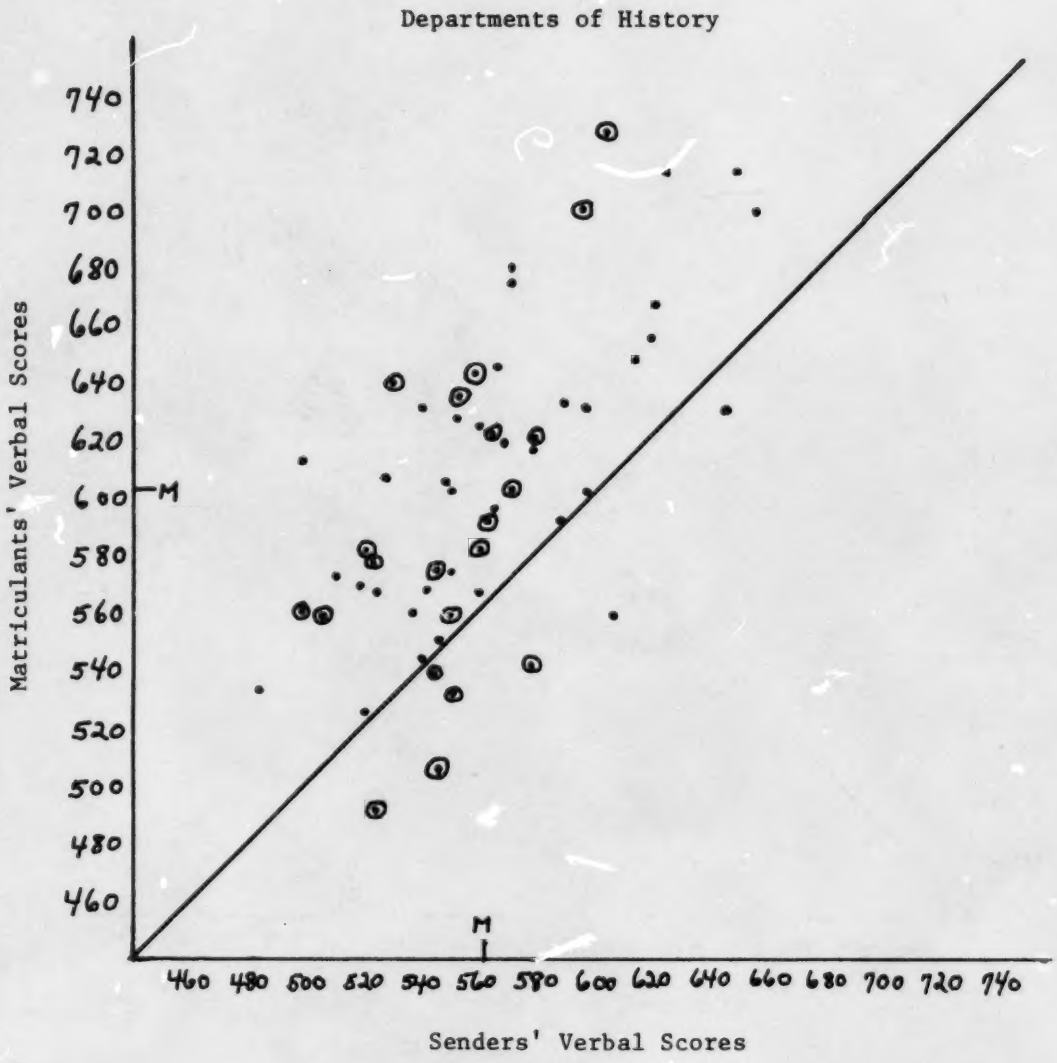

Fig. 1. Scatterplot of mean enrolled vs. mean sender GRE verbal scores in history departments.

Note: Circled points indicate departments with matriculant scores based on fewer than 10 enrolled students. The diagonal line represents perfect agreement between the two sets of scores.

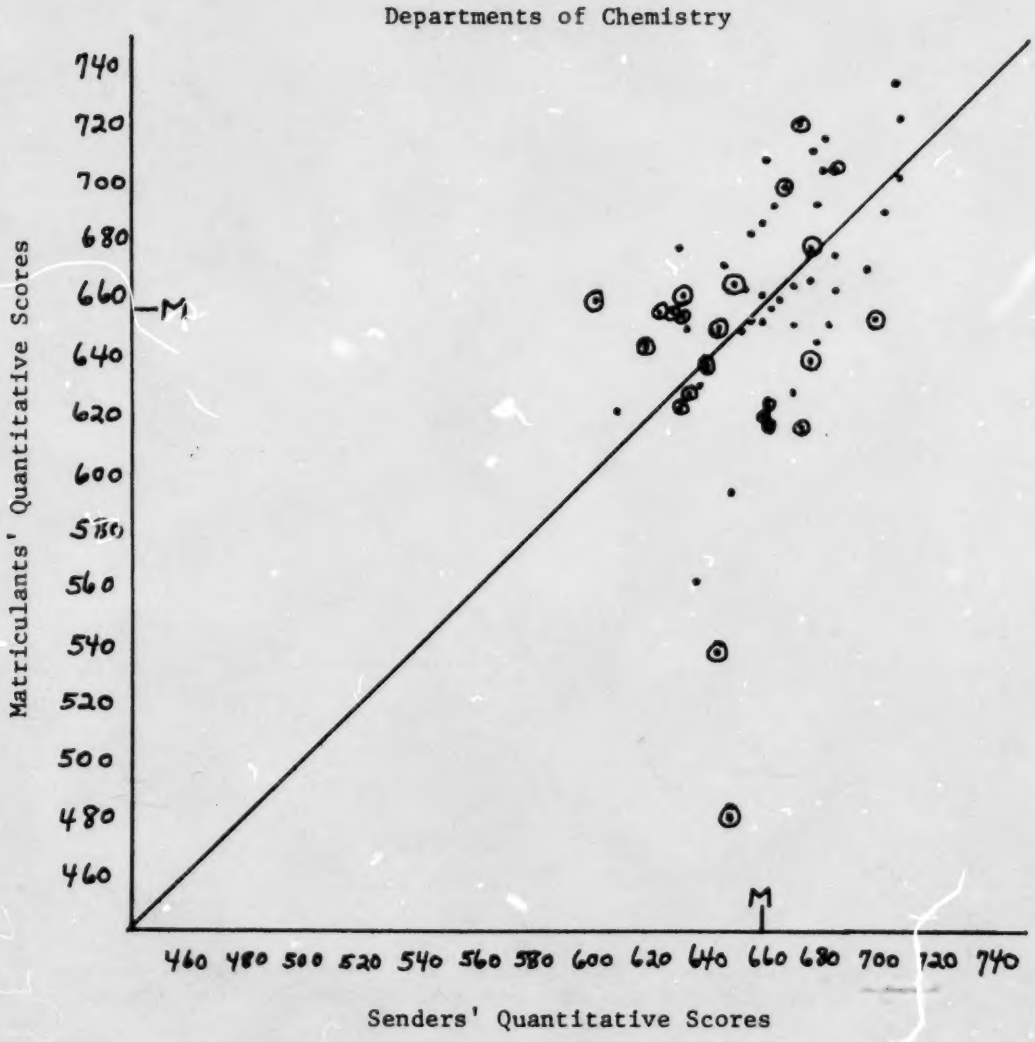

Figure 2. Scatterplot of mean enrolled vs. mean sender GRE quantitative scores in chemistry departments.

Note: Circled points indicate departments with matriculant scores based on fewer than 10 enrolled students. The diagonal line represents perfect agreement between the two sets of scores.

in history, with high scoring test-takers sending their scores to a larger number of "safety" departments. There is some slight evidence to support this interpretation when sender vs. matriculant scores are examined in relation to departmental reputational ratings, but only in the departments with middle-level reputations (Roose-Andersen ratings between 3.49 and 2.0). As shown in Figure 3, the average quantitative GRE scores are higher in the applicant pool than among enrolled students for more than half of the moderately rated departments (18 out of 32). Only among the lowest rated programs do more than half of the programs enroll students with a mean GRE-Q score that is higher than the mean GRE-Q score of prospective applicants (seven with matriculant scores higher than sender scores, five the reverse).

The relative status of mean enrolled vs. sender scores on the GRE-verbal test for historians also indicates that the departments with low peer ratings are more likely to enroll higher-scoring students relative to their applicant pool than are the top-rated programs. That is, when the departments are rank-ordered on the mean verbal sender and matriculant scores and the differences in rank-orders are examined, as in Figure 4, more than half of the low-rated departments enroll students with verbal scores that improve their rank (8 vs. 3) while more than half of the top-rated programs (4 vs. 2) enroll students with average verbal scores that rank the department slightly lower than its rank based on the verbal scores of prospective applicants. Departments with mid-range ratings are evenly distributed in relative rank-orders based on average enrolled vs. sender scores.

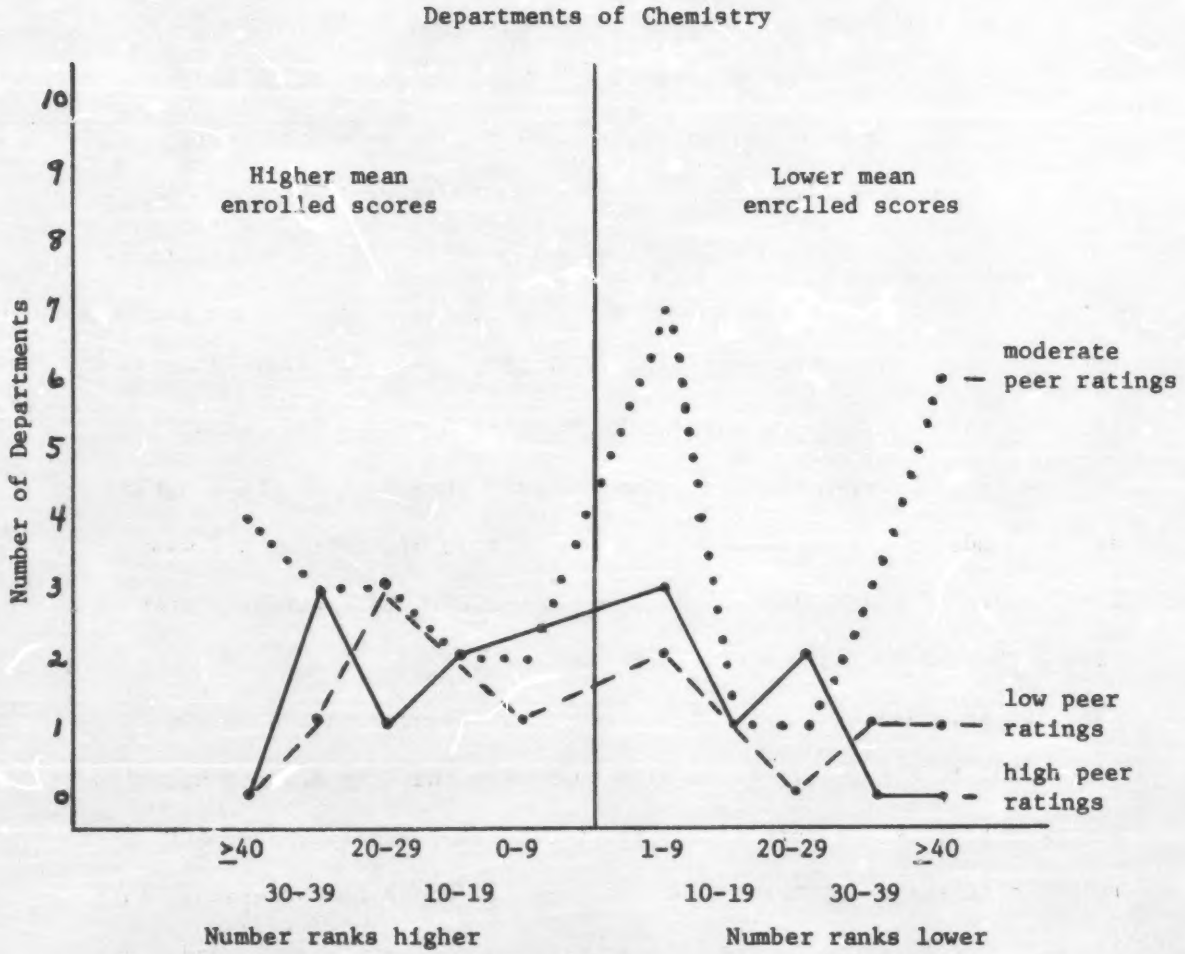

Figure 3. Differential rank orders of mean enrolled vs. mean sender GRE quantitative scores in 56 chemistry departments with five or more GRE scores for enrolled students, by reputational ratings of the departments.

——— = Roose-Andersen ratings 3.5 or above (outstanding)

• • • • • = Roose-Andersen ratings 2.0-3.49 (adequate or better)

— — — = Roose-Anderson ratings below 2.0 (less than adequate)

Peer rating/sender scores r = .74

Peer rating/enrolled scores r = .63

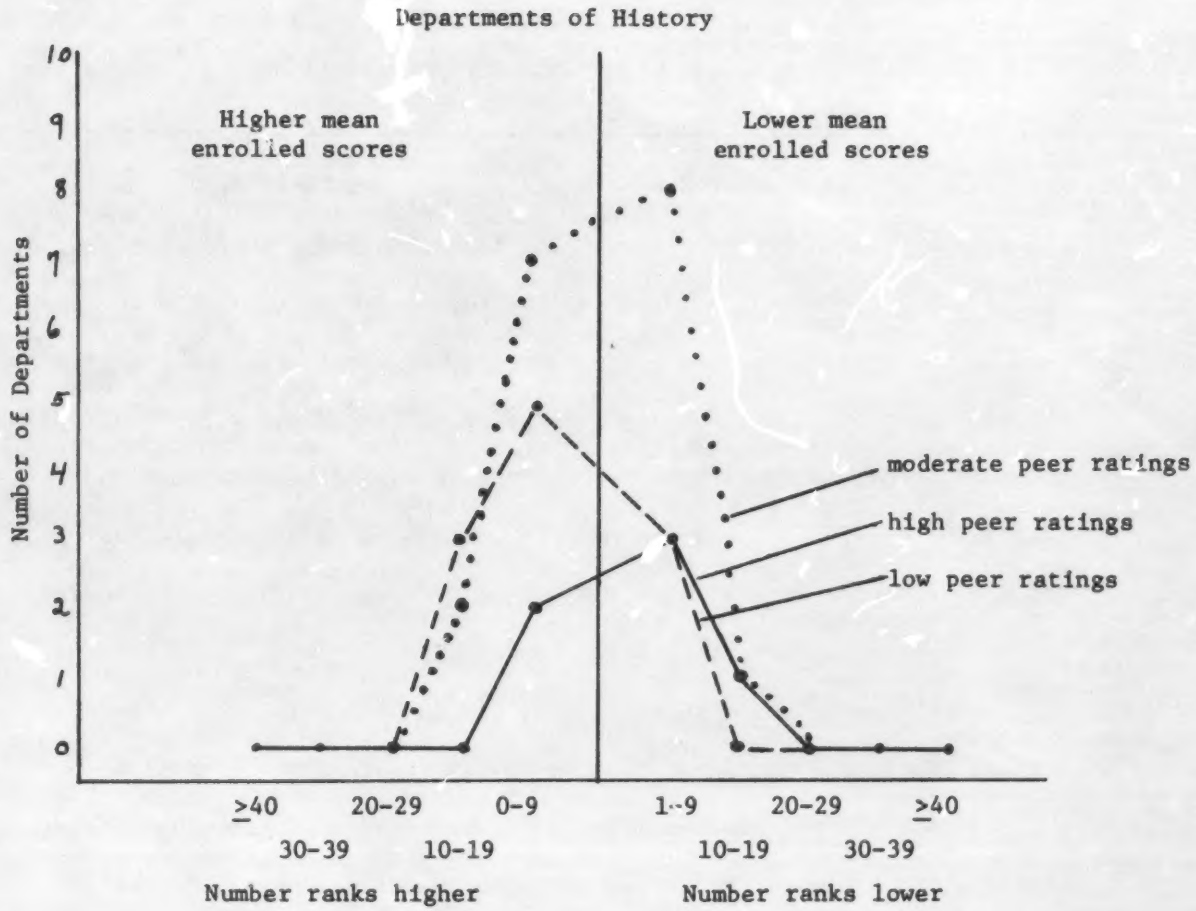

Figure 4. Differential rank orders of mean enrolled vs. mean sender GRE verbal scores in 37 history departments with five or more GRE scores for enrolled students, by reputational ratings of the departments.

——— = Roose-Andersen ratings 3.5 or above (outstanding)

• • • • • = Roose-Andersen ratings 2.0-3.49 (adequate or better)

— — — = Roose-Anderson ratings below 2.0 (less than adequate)

Peer rating/sender scores r = .90

Peer rating/enrolled scores r = .65

The data examined so far suggest that much of the admissions selectivity often ascribed to graduate departments might more accurately be labeled student selectivity in deciding where to apply. Student self-selection is further demonstrated by a higher level of agreement between the reputation of a department (as represented by the peer ratings of professional colleagues) and the mean GRE scores sent by prospective applicants than between such reputational ratings and the GRE scores of enrolled students. As reported with Figures 3 and 4, there is a correlation of .90 between the Roose-Andersen reputational ratings of departments and the mean verbal scores of senders in history, and a correlation of .74 between the mean quantitative scores of senders and the reputational ratings of chemistry departments.[3]

The number of departments in both fields with relatively higher average GRE scores among prospective applicants than among enrolled students was not anticipated. The more prestigious programs enroll more high-scoring students, as reflected in correlations of .65 between reputational ratings and average matriculant verbal scores in history and .63 between ratings and average matriculant quantitative scores in chemistry, but these correlations are lower than the relationships between reputational ratings and sender scores. Such findings suggest that a traditional measure of departmental admissions selectivity such as the ratio of applicants to admissions is not likely to be appropriate in

---

[3]These correlations are based on departments in the sample with five or more test scores for matriculated students, and on mean ratings of the quality of the graduate faculty (scale 5-0) for individual departments that were obtained from the American Council on Education.

academically oriented graduate departments, since more sorting takes place in the application than in the selection stage of the process. Also, the data suggest that many factors in addition to test scores are being taken into account in the admissions process. Given a sizeable pool of applicants reasonably well matched relative to the academic demands of a program, these results support the frequent claims of departments that they do not always admit applicants simply because they have high test scores. In general, it appears that a graduate program in chemistry or history, regardless of reputational rating, has little more than a 50-50 chance of enrolling graduate students with relatively higher academic ability than is represented in its applicant pool.

Returning to the original question of whether or not the departmental average of GRE sender scores might be used as a proxy for departmental selectivity, the evidence is not clear-cut. Data from a fairly large number of departments in two disciplines indicate that there is a positive relationship between sender and matriculant scores, with correlations in the range of .5 to .7. Correlation coefficients of this magnitude are quite useful in an admissions context where the task is to select individual students. However, it should be remembered that these are correlations between department means rather than the scores of individual test-takers; correlations in this range that are based on pre-aggregated data indicate a fairly wide margin for error when making judgments about departments. It is doubtful that correlations of this size justify the substitution of mean sender scores for the mean scores of enrolled graduate students in a particular department, though the sender scores might be

useful in certain kinds of large-scale research where comparisons between individual departments were not being made.

Some appreciation for the likely magnitude of the errors when one mean is used to estimate the other can be gained by looking again at Figures 1 and 2. In Figure 1, a history department mean verbal score of about 520 for senders is associated with mean matriculant scores ranging from about 490 to 580; a mean verbal score of 610 for senders is associated with mean matriculant scores of about 560 to 650. In chemistry (Figure 2), a mean quantitative score of about 650 for senders is associated with mean matriculant scores from about 480 to 680 (560 to 680 if the outliers are omitted) and a mean quantitative score of 700 for senders is associated with mean matriculant scores from about 650 to 740. Omitting departments with small enrollments does not reduce these ranges appreciably. Most persons associated with the graduate admissions process would consider these to be quite important differences, particularly if the sender scores were substituted for matriculant scores in materials that were intended to describe a department. Obviously, departmental selection procedures and student enrollment decisions include factors that are not well represented by the mean GRE scores of potential program applicants.

From these perspectives, we conclude that average GRE sender scores are not very accurate proxies for the average test scores of students enrolling in individual graduate departments, despite a positive association between mean sender and matriculant scores.